JPRS: 3853

29 September 1960

FOREIGN DEVELOPMENTS IN INFORMATION PROCESSING

AND

MACHINE TRANSLATION

No. 1

RETURN TO MAIN FILE

19990611 130

FOREWORD

This publication was prepared under contract by the UNITED STATES JOINT PUBLICATIONS RESEARCH SERVICE, a federal government organization established to service the translation and research needs of the various government departments.

FOREIGN DEVELOPMENTS IN INFORMATION PROCESSING
AND
MACHINE TRANSLATION
No. 1

FOREWORD

This translation series presents information from Soviet literature
on developments in the following fields in information processing and
machine translation: organization, storage and retrieval of information;
coding; programming; character and pattern recognition; logical design
of information and translation machines; linguistic analysis with
machine translation application; mathematical and applied linguistics;
machine translation studies. The series is published as an aid to
U. S. Government research.

This is the first report of this series.

EXISTING LAWS RELATING TO COPYRIGHT MAKE
NECESSARY THAT THIS TEXT BE RESTRICTED .
TO THE OFFICIAL USE OF THE U. S. GOVERNMENT

MACHINE TRANSLATION AND APPLIED LINGUISTICS

Applied Linguistics, a science making rapid strides.

It appears more and more clearly today that the application of
automation to the mechanisms of thought is directing us toward a
revolution in the means of communication and in the conservation of
the explored and conscious content of the human mind, comparable to
the revolution in the art of printing in the sixteenth century. For
all those whose professional and intellectual life revolves around the
written or the spoken word, or whose communications with their con-
temporaries or with the future depend on writing, this twentieth
century revolution calls for a re-examination of the methods and the
aims of their communications. For those in short who in different
degrees and for different purposes are interested in the study of
language, a new constellation of disciplines is taking place within
the confines of cybernetics and the new linguistics: its content and
its frontiers vary, moreover, according to country and school but in
Moscow as in Washington it is given the same name -- applied linguistics.

Applied sciences have their origin chiefly in the needs of the
society in which they develop. Unquestionably there are few examples
of a pure science alone leading to the birth of an applied science,
without the impetus of a social need. In our times it is need which
permits us to release the financial means indispensable to research
which is somewhat complex and thus devours time and talent. This is
certainly the case in applied linguistics. Neither scientific curiosity
nor the ingenuity of phoneticians have done more for the study of certain
laws of language which are susceptible of practical applications than,
for example, research on better conditions of telephonic transmissions;
studies by telecommunications engineers on transmission of signals have
resulted in the information theory in which linguistics finds new ideas
and new means for the unbiased and even philosophical study of language.
These are only beginnings.

Multiple and diverse in effect, the needs of our time which
oblige us to know better the fundamental laws of language corresponding
to the growing complexity of relations between human groups on the
surface of our planet, soon perhaps in inter-astral space. The expansion
of these needs coincides furthermore with the maturation of peoples
having recently gained independence, with the multiplication of inter-
national relations, the creator of immense needs for translation; with

the expansion of our knowledge, requiring a growing number of words to designate it; as also with the necessity, in which man finds himself face to face with his own inventions, to give to his language a rigorous exactness in expression, because this language must also serve more and more to establish a communication between the human brain and machines. The latter, which do not think, being like artificial bodies, muscles and nerves which are added to the human body, must receive signals devoid of ambiguity, either in order to act instantaneously or to be able to correctly interpret ambiguous signals, thanks to their logical working which is certainly admirable, but limited all the same. The machine invites us to "debabelize" ourselves; now Babel is everywhere, even in those places where people have the illusion of speaking the same language. Anti-babel is a necessity of our civilization; language is a tool whose precision must grow if our civilization wishes to preserve its cohesion and to be able to be spread by ways of intelligence and not by a more or less instinctive traditionalism.

Synchronic linguistics and the study of signals.
           Around the system of conventional signals which all language is, there has been organized a whole gamut of research illuminating each one of the multiple and mutually tangled aspects of communication systems belonging to man, going from the first cries uttered by the new born baby to attract his mother's attention, to the most complex expressions of tribal and family relationships, to abstract structures of the loftiest philosophy, to the euphemisms or lies of the lowest politics. And whether we wish it or not, meanings, from that of the black sail of Theseus to that of the Kantian "Vernunft" [reason], from that of the blinking beacon, the Morse radio signal, the first syllable stammered by the child, to Einstein's formula $E = mc^2$, will be the legitimate subject of study of applied linguistics, as will be all the different aspects of the relationships between the signifying and the signified. The question what does it mean? asked by the anxious mother trying to understand the cries of her baby, by the radio which has just picked up an unintelligible signal, by Champollion in front of the Rosetta stone, by the translator faced with a sentence containing a hermetically sealed message either in itself or because it is not in the intellectual baggage of this translator, this question lies at the center of the problems of all applied linguistics, it is the raison d'etre of this science.
           But it is no answer to this why if we do not have first of all the answer to innumerable how. Around this central problem of meaning, to which all the others lead us back, researches closely connected with this problem as they are among themselves are being performed in stages and organized. All the resources of pure linguistics come into play except perhaps for its historical resources: linguistic studies of our times are put under the sign of synchrony, contrary to the diachronic linguistics of a short time ago. The passage of language signals in machines poses all the questions about materializations of thought by

sounds and visual signals. Phonemes and graphemes and their complex relations must be the object of rigorous studies. Likewise with morphemes and tagmemes, sememes and their meanings in the present. For it is the whole complex system of the meaning of language which is involved; each separate study touches another neighboring study but which must remain distinct so that analysis may progress. Even stylistic studies will belong to applied linguistics. Although they must bear on some aspects not strictly cognitive of language, it will be essential to map out clearly the boundary between these cognitive aspects and the neighboring and sometimes overlapping domains peculiar to the aesthetics of the verb. If language is in effect a plastic substance moulded at their will by artists -- poets or prose-writers -- it is on a double game bearing on the cognitive and the aesthetic that their art is founded, in which as Roman Jakobson recalls, the play on words or paronomasia [misleading similarity between words belonging to different languages] holds sway.

Privileged role of translation.

If translation has again become the question of the moment it is among other reasons because it lies at the very heart of this crux of problems, never yet completely separated by a definite analysis. Now researches undertaken with a view to the automation of the translation of scientific and technical texts permit and require us to isolate in discourse the diverse elements of the transmitted message -- phonic, graphic, lexico-semiotic, aesthetic, etc., elements. They demand, at least at first, that we study the graphemes independently of the phonemes and that we take the written language as a basis -- even if this effort must stress the inconsistencies of the alphabetic trans-criptions of certain languages, written for too long a time for the simplicity of their orthography. Then, beyond these distinctions made at the level of the constituent signals of the message there are others which become necessary and possible: works with a view to machine translation have started analyses and experiments going which permit us to isolate cognitive elements of discourse, stylistic or aesthetic elements; we shall then have to distinguish between signals, the cogni-tive meaning, and these additional aspects of meaning which are within the province of aesthetic suggestion rather than of information. The methods of analytic experimentation required by the automation of language bring to linguistics the means of scientific investigation permitting a more rigorous separation between information and poetry in the analysis of written messages. Likewise they require a rigorous application of logical analysis for the determination of the content of the authentic information of a message.

There are other profound reasons which give a special role to machine translation at the heart of problems of applied linguistics. Man has never invented a richer or more varied system of expression than language. Plastic artists and musicians will pardon me if I stress the fact that the semiotic value of their works can be described by words,

more often and more faithfully than their art by its own means can evoke all that language expresses. And the semiotic system which language is is essentially cognitive and descriptive to the extent that it is above all social. For the actual environment of man language is a mirror, it determines the abstract environment and the structures of the mind, the dynamism of gradual development; it also knows how to espouse and evoke this dynamism, as music can. This universe of experience and the universe of discourse tend to be equated and to be confused. To try to mechanize the transposition of the written or spoken expression of a language into another is thus to try to establish a double equation between the universe of experience, and two universes of discourse which reflect it, but with different iridescences. To partially succeed is to affirm and demonstrate the function of language as a tool of our human civilization, and to spread its usefulness and to increase its value as a means of understanding among men.

Once the machine is capable of translating sentences from one language into another, and is endowed with a memory sufficient to do it on a large scale, it will also be capable of retrieving and classifying concepts, that is, to retrieve information without passing through artificial and more or less arbitrary codings as we search for it almost everywhere at the present time for the needs of documentation. The translation machine will doubtless bring a fertile means of analysis and classification of knowledge, having as a base the words of the dictionary. In translation, whether machine or human, the word or the semantic unit evokes the idea which in its turn evokes the word of the target language. This mechanism presupposes a well-ordered memory, which can be useful in classifying library titles and in classifying the analytics of works indexed in a bibliography. The universe of well classified discourse will thus serve as a guide to the reader through the shelves of a library. Translation directs us towards complete solutions to the problem of quest for information.

Programs of machine translation improved by a cumulative process.

Where is machine translation today? As for what has been done from 1946 to 1959, permit me to refer the reader to my recent book (E. Delavenay -- La Machine a traduire -- Paris, Presses Universitaires de France, 1959. Introduction to machine translation. London, Thames & Hudson, 1960), to a lecture published in the magazine Impact (Impact, Science et Societe, Vol 10, No 1), to two articles by Michael Corbe in the magazine Automatisme (Automatisme [Automation], Vol 4 (1959), Nos 7-8 and No 12) discussing the particulars of Russian and American work, and to the studies of Mr. A. Sestier in Ingenieurs et techniciens [Engineers and technicians] (Sestier (A.) -- La Traduction automatique des textes ecrits scientifiques et techniques d'un langage dans un autre [Machine translation of scientific and technical written texts from one language into another] in Ingenieurs et techniciens, March,

April, May and June 1959). Although the accent is put by some on syntax or structures, by others on the lexicon, some partial translations today are leaving machines which have not been specially constructed for this usage. These translations are reviewed and refined by "editors" who inscribe in the memory of the machine the semantic equivalents and the rules of syntax or of translation which comparison of the original and its translation reveals the necessity for. Thus a course of progressive and cumulative improvement of programs is pursued. However, much depends, and will more and more depend, on the basic concepts of these programs. Such a system of translation which gives rather encouraging results at first sight can quite easily become inoperative and as if paralyzed by the accumulation of its own refinements. In fact each new added rule could, by reason of an error in the initial conception or through insufficient analysis of the problem of translation, become a cause of a slowing down in the execution of the program. So much that the perfect program with which we could theoretically end up would cost more than the best human translator. If machine translations which are imperfect but immediately usable by specialists in the subject in question are possible from now on, fully automatic and high quality translation, to go back to Bar Hillel's expression, is nowhere to be found today and will doubtless only reach that level, thanks to basic research bearing on the deep foundations of language as a semiotic system.

The double road of discovery.

Thus a double route is outlined for the discovery: on the one hand we shall have to pursue and go thoroughly into more or less empirical studies of bilingual lexicographical analysis and of comparative structural analysis which have made possible the present day work of American teams like those of the Rand Corporation, Ramo-Wooldridge, Seattle, Harvard, and Georgetown; or of Russian teams like those of the Steklov Institute of mathematics and the Institute of numerical calculus and precision mechanics. Without such studies we could not hope to achieve the storage in the memories of electronic machines of quantities of linguistic facts sufficient to permit us to reach the next stage. However, it is, on the other hand, indispensible to be able to view the entirety of the researches in applied linguistics from a synthetic viewpoint, and to put researchers to work who are capable of coming up with a general theory of linguistic meaning, in order to lay the foundations for a universal program of translation. It is difficult to dispute the long-term utility of such research and we can only really discuss their being spread out over time and their theoretical points of departure. But it is not to diminish the important role of logic and mathematics in a totality of works which language remains the object of, as to emphasize that the theoretical positions of these sciences will be without a basis as long as there does not exist the sum total of precise observations of linguistic facts on which they will probably be able to operate. We must translate

by machine, as quickly as possible and as much as possible, in order
to check the facts which will permit us to do it better and perhaps
to discover universal programs more economical than the bilateral
programs.

But which road to follow in order to arrive at this point of
departure! And how many choices are given from the outset, each one
limiting the possibilities of subsequent choices and conditioning the
solutions of the entirety of the problems.

Training men to use machines.

We should not let ourselves be unduly influenced by considerations
relating to the type of machine which will eventually be used for trans-
lation. There exist today numerous electronic calculating machines
with different characteristics. The fact that the translation machine
will probably be a machine endowed with a highly developed memory
(with a large capacity and with ultra-rapid access) and with relatively
less powerful logical apparatus, will not prevent us from temporarily
using existing machines made to solve administrative or scientific
problems of quite another nature than those of translation. Above all
it will be essential to assure ourselves that the machines employed
are a part of assemblies adaptable to various purposes and whose parts
can be combined at will to answer these purposes — a preoccupation
which moreover coincides with that of great designers who are anxious
to extend the applications of their equipment. Above all, it will be
necessary to utilize the equipment widely and have it easily available
so as to be able to develop to the maximum in accordance with an over-
all plan the preliminary projects, those in which close collaboration
of man and machine requires the use of machines which can wait for
their slower partner.

The number of hours of patient preparation devoted to the
compilation of bilingual dictionaries and statements of structures,
and to putting them into shape for automatic processing by machines,
will be considerable in comparison with the time required to make a
translation of a few hundred lines on an IBM 704 or on a Gamma 60.
It will be possible to work on sections of programs, in separate
teams, first having access to card punching machines, then to cal-
culating machines, on condition that the general and technical data
of the problem have been defined in advance and that each team con-
forms strictly to the general plan, which can very probably be con-
ceived for any electronic calculating machine rather than for any
given model of calculating machine. In any case it will be essential
to permit linguists to work free of any technical preoccupation related
to machine programs or to the type of machine itself, in a setting in
which they can advantageously apply their knowledge in preparing the
language for machine processing. For a very large part of the
preparatory work, neither the type of machine used nor the type of
program conceived for this machine will be of any importance whatever
in the analysis of the language.

Memory and logic, two aspects of language prepared for the machine.

On the other hand it is certain that the machine will have a memory and a logical apparatus permitting it not only to search for words and translation rules in this memory, but also to modify the order and the nature of its sub-programs in accordance with the indications found in the memory. The relative importance of data put into the memory in a static manner, and of programs and sub-programs of analysis and synthesis, can vary, just as the methods according to which the indications furnished by linguists on the input language and the output language will be ordered in the various parts of the machine. Definitive decisions on those points will depend on the type of machine adopted and in large measure on the vocabulary and on the structure of the two languages on which we shall work, as on their morphology, in short on what Andreev calls their "congruence."

One of the first tasks consists in enumerating all the indications before being noted down in the memory or memories in connection with each word, in classifying them as to whether they are constant, whatever the function or the form of the word in the phrase may be, or whether peculiar to this word in a given context. This distinction between constant and contextual information relative to the word will ultimately have to allow us to divide up the tasks between static memory and dynamic programs, the stable lexical characteristics of the word probably being inscribed in the lexical memory, while its contextual and therefore unstable characteristics will be the object of instructions for more or less detailed sub-programs.

This preliminary work can be carried out by teams with relatively little knowledge of the requirements of the machine, provided they are well experienced, and that their works are integrated in a preconceived plan and are submitted to checks and to regular improvements in the light of the first tests permitting the testing of notations and programs which they are setting in motion. Within these teams the presence of programming technicians for electronic calculating machines is indispensable in order that the linguists may not lose sight of the characteristics peculiar to the type of analysis that will be required of them: an analysis made for a machine without a flicker of conscience, to which everything must be told because it never guesses.

Beginning by enumerations.

Just as in the case of machines, there is no need at all to prematurely take a position on various linguistic theories or skeletons of such theories. It is very tempting to think that a general explanation of the facts of language should greatly accelerate research with a view to automation of translation and information research. However, language reflects the universe of representations, concrete or abstract, particular or general; and the information communicated by it varies in particulars or in generalities as the intentions of the

subject speaking. This is why translation, which is a transposition
of information -- quite as much as information research, which is
based on the detailed or analytic recording of researched facts --
cannot without loss of information deviate much from the degree of
detail desired by the author of a text. It is therefore very difficult
to imagine how a general theory which is other than a general taxonomy,
a methodical inventory, will ever permit an important simplification
of a problem at the base of which we always find the word, that is,
either the grammatical tool or the signifying element establishing
a relationship between a mind and a signified element.

This is why we judge that between the empiricists of machine
translation and the theoreticians seeking a general explanation, there
is in the final analysis opposition only on slight differences, some
being more conscious of the unstable diversity of language, others
more desirous of being able to generalize the rather rigerous laws
which govern all language insofar as it is a systematic code and
social convention. This is also why the options between lexical
empiricism, structuralism, mathematical formalization, etc., are in
short rather unreal, and it is quite necessary to reach the point of
this "combined theoretical and empirical line of attack" defined by
Reifler and which has given excellent results. (Linguistic and
engineering studies in the automatic translation of scientific Russian
into English -- The University of Washington, 1959. Outline of the
Project: Intermediate MT systems as research tools.) It will be a
matter of enumerating words, the meaning of these words, the rules of
association of the words among themselves, rather than to search for
a philosopher's stone for all translations.

The teams exploring the relationships between two given languages
with a view to bilateral unidirectional machine translation (we shall
already have made an important choice in thus defining in a restrictive
manner the role of these teams and in excluding at the outset a general
research in universal translation) will nevertheless have to make every
effort to work in such a way that no option important to a subsequent
stage of research is excluded. It is in fact necessary to be able to
take maximum advantage of their work at any moment, whatever the option
chosen at that moment in matters of machine or program for the sequence
of operations. These teams will have to go progressively from the
particular to the general, the particular really only being able at
the outset to be a corpus or collection of texts of the input language,
in which each word will be studied in the double connection of its
constant and contextual characteristics, and naturally as well from
the viewpoint or morphology if it is a matter of a language which is
more or less inflected. However, besides this lexicographical work
they will be able simultaneously to begin the study of its structural
aspect; for example, a group will be interested in making an inventory
of the words and in the identification of their grammatical char-
acteristics, while another will study structures in the same corpus --
a study closely connected to that of the constant and contextual gram-
matical characteristics of words, and capable of also having ties with
their examination from the semantic viewpoint.

- 8 -

We see here that a third study will be necessary: that of the meaning of words of the input language from the viewpoint of their translation into the output language. This semantic analysis is done from the angle of correspondences in meaning from language to language and in the comparative examination of the area of meaning of a word of language A and of the word having the same meaning in language B. Since these three studies can be carried on by three different teams, it will be necessary to note, in accordance with common standards, all interferences of each of three studies on each other, which will lead quite rapidly to isolation and classification of the most complex problems, that is to say those in which structure, morphology, lexicon and probable meaning would have to come into play to resolve an ambiguity. It is this study of interferences which will result in the formulation of translation programs of difficult structures. As for the easy cases, the inventory of words and structures by itself will have permitted us to resolve them by simple equivalence from language to language. It will be advisable to determine as soon as possible the proportion of simple cases and difficult cases for a pair of languages.

Coding, symbolizations, auto-programming.

The works sketched above being properly linguistic and able to be entrusted to men having only elementary knowledge of machines and their functioning, they will not however be conceived independently of a work project on machines which must begin as soon as possible with a view to the gradual mechanization of the greatest part of the research. In fact each act of inventorying forms, words, structures and meanings must result in a brief, and as much as possible, coded, notation which can be used directly, or transcoded either on punched cards or on any other support accessible to an electronic machine for the purposes of statistical work, classifications and translation. A first coherent and systematic coding of linguistic notations from the "human" stage of the work will be necessary if we wish to mechanize the projects from the outset as soon as possible; it is probable that trial and error attempts will be necessary at first, even granting the most complete and perfect cooperation between calculating machine engineers and linguists. At a relatively advanced level of research a first analysis of the linguistic and electronic data of the problem, conceived so as to leave open the possibilities of solving unforeseen problems, must be made in order that the general framework of coding and of mathematical symbolization of the translation programs may be mapped out, in which the observations of linguists will have to be integrated. All proportions will be kept the equivalent of the COMIT of Yngve and M.I.T.

It is particularly in the realm of automatic recognition of structures that collaboration from the first between linguists and engineers will be necessary and productive: it is actually necessary

not only to distinguish and classify the structures of elements in the phrase, but to connect them to symbols expressing the possible functions of words, to establish fixed relationships between these symbols and certain structures, to express the whole by a rigorous and as economical as possible symbolization, and, for certain cases of ambiguity, to conceive a program permitting the machine to try several structures and to choose from among them the one which it is possible "to prove."

Once these stages are gotten over, it will be possible to begin different types of rudimentary machine translations. Some will be word-for-word, structurally deficient because they will be based on a simply lexical analysis, and will leave to the reader the choice between various constructions and multiple meanings. Others, on the basis of structural analysis, will be "block-for-block" in which we shall recognize sequences of words correctly ordered, many words not yet however being translated, and appearing in the input language; others will result from the combination of the two programs and from progressive improvement, thanks to the critical examination which it will be the object of; little by little rough drafts of more perfected translations which are more and more comprehensible to others than machine translation technicians will result.

Sketch of a program.

This outline of work which will allow us to begin to translate by machine —— that is, to put ourselves in a position to discover the means of improving machine translations, to utilize the inventories of facts effected during their preparation, and the critique of the programs utilized —— is naturally to suggest broad lines of action which can advantageously be undertaken where it has not yet been done, notably in France, with a view to the development of applied linguistics. It is very evident that research must be carried out simultaneously on various planes —— that of technics and industry on the one hand, that of linguistics on the other. And whoever says linguistic says social since all concerted action on language takes place in discussions and in reactions as soon as it collides head on or even indirectly with certain unconscious habits. We know, for example, what difficulties arise in the face of the most timid attempts to normalize trans-literations and how much each one holds on to totally anarchic ortho-graphic habits. Because we will not be able to act on all fronts at once, and because the technics front will proceed much faster than the social, it is therefore necessary to establish an overall plan, to choose the regions for priority action, to keep contact with all aspects of the problem without hoping to be able to act simultaneously every-where.

Care in developing and improving existing machines can be left to the designers and their research bureaus. Here needs are known long in advance in most cases and research is in progress. Improvement in the capacity and speed of access of the memories, implementation of

machine programs which can discover the solution to set problems without human intervention, creation of apparatuses for automatic reading of printed texts, studies of apparatuses for alphabetically transcribing spoken words, all that is in progress; it will suffice for linguistics to become conscious of their role, to define in terms understandable to engineers the problems which are peculiar to them. There is a primary and urgent need in which a concerted interdisciplinary action is necessary.

A new scientific association.

It is therefore necessary first of all to train men for this interdisciplinary action and for that purpose to diffuse ideas and knowledge about applied linguistics whose development is accelerating. Has not a Center for applied linguistics, founded by the Modern Language Association of America, just been set up in Washington with the aid of powerful foundations? In Leningrad, N. D. Andreev emphatically defined the tasks of this science (Andreev (N.D.) and Zinder (L.R.) — Fundamental problems of applied linguistics, in Voprosy Jazykoznanija [Problems of Linguistics], 1959, No 4), while in Moscow an Institute of applied linguistics has been founded and a magazine entitled Traduction mechanique et linguistique appliquee [Machine translation and applied linguistics] is being published. (Mashinnyi perevod i prikldnaja lingvistika [Machine translation and applied linguistics], 1959.) To expand and thoroughly explore the knowledge in this field, to develop the attitude of scientific observation in the face of linguistic fact, to inventory problems, to seek connections between themselves and with related sciences and disciplines, these are some of the primary tasks to accomplish.

Because of the feeling of the urgency of these tasks and the desire to encourage in France the implementation of the means suitable for accomplishing it, there was born in Paris the Association pour l'Etude et le Developpement de la Traduction automatique et de la Linguistique appliquee (ATALA) [Association for the Study and Development of Machine Translation and Applied Linguistics], founded in the Fall of 1959, with the aim of informing and training men by the exchange of their knowledge. It is a continuation of a small international group for studies on machine translation which the author of these lines initiated in 1958 with a few colleagues; the reports presented by the members of the group before several invited guests permitted a rather rapid assimilation of certain interdisciplinary aspects of the research, and convinced each one of the advantage of these meetings; soon a sub-group was attacking problems of machine translation from English to French. Since the summer of 1959, the decision was made to transform the Group which had grown rapidly by co-option into a French Association conforming to the law of 1901, as a non-profit institution and dedicated to educational and scientific aims. It was accomplished as of October 1959 and the Association immediately started work to study its means of action and to define its scientific program.

- 11 -

The aims of the ATALA are thus defined in its statutes:
"Article 2.

"The Association has for its aims the encouragement of the study of linguistic, scientific and technical problems of machine translation as well as the coordination and organization of individual efforts in this field, and to thus facilitate scientific and cultural exchanges of a kind calculated to strengthen peace. To achieve these aims it proposes in particular:

a) to assure that its members are mutually informed on machine translation and on fields and connected techniques, notably by the publication of a periodical or bulletin, works and all other documents concerning the field of its activity;

b) to establish, with a view to exchanging information, contacts with all other groups interested in the same subjects as itself;

c) to encourage and to aid its members to carry out practical work or special studies in the above fields;

d) to participate actively in the diffusion of knowledge acquired in the electronic processing of linguistic and logical data;

e) more generally to contribute, in all useful forms and by all appropriate means, to the realization of machine translation and to the development of applied linguistics.

"Article 3.

"The Association will strive to promote the creation of an International Federation of national groups having the same aim and will eventually be able to set up the French section of this Federation."

The Association is engaged in giving itself as precise a program as possible, for the elaboration of which the following principles obtain: to preserve for the Association its character of a free meeting place of all good wills, and its independence; to avoid duplication with all existing or future official or private institutions carrying on a connected activity, and to seek, on the contrary, any occasion to help this activity by means befitting a free association; to concern itself with facilitating on the national and international plane the most complete exchanges of information in order to stimulate research and the diffusion of knowledge; to undertake as soon as possible research pilot-projects, which would only be to keep its members in close contact with the reality of practical work, and never to confine itself in an academic and theoretical atmosphere.

The implementation of this program will depend to a certain extent on material support by which the Association will profit --supplying services by large firms specializing in the manufacture and use of electronic machines, exceptional contributions of member benefactors or research contracts and subsidies.

Pending the time these plans are made specific, a primary, provisional program comprises the following points: 1) Work of a practical order on machine translation; 2) formation of cadres; 3) communications and information; 4) material organization of the work; 5) theoretical work.

Practical works: it has been decided after a big discussion that these works would at the outset bear on translation from English into French, which has not been undertaken or studied up to now by any other group. The first efforts will consist in defining a methodology. Already a scientific committee of the Association is at work. It has presented a highly encouraging first report, has divided up the preliminary tasks and has inventoried the problems.

The possibility of undertaking work on other languages has not been excluded but it will have to wait until the number of active members is sufficient to permit a large enough group to take it up. It will be interesting for example to study the characteristics of translation from one Romance language to another in order to determine the nature and importance of the difficulties encountered, and to search for the possibilities of "common parent stock" in translation programs between languages of the same family.

Instruction. The regular scientific meetings of the Association and of its working Committees already constitute one of the most valuable forms of advanced instruction and continue the work of the former international study group. The initial recruitment of new members in university and industrial environments has exceeded the hopes of the first founders, and the Association is in a position to unite very diverse competences and to illuminate numerous problems of applied linguistics, in calling on knowledge of its members. When the need arises it will be able to furnish any educational institution making a request with lecturers capable of dealing with general or particular subjects. It is stipulated moreover to formalize as soon as possible this mutual instruction and to extend it, for example by organizing a regular seminar in connection with an institution of higher learning. It could have as its object a critical historical study of what has been done in the field of machine translation; it could also broach the study of a problem at once extended and precise, as for example that of the pivot language or universal language of the machine, closely connected with the topic of the formalization of natural languages with a view to documentary research.

The work of the scientific committees of the Association will permit us to determine in the next few months to what extent the latter will be ready during the second half of 1960 to undertake these tasks in connection with a public institution of advanced studies.

Communications and information.
As much for the internal relations of the Association, of which all members cannot participate in all projects, as for its external relations in France, and abroad where it has already been assured of valuable contacts, the creation of a liaison organ permitting the exchange of ideas and information is indispensable. A committee has been charged with preparing the publication of this bulletin, which will be first of all a stenciled sheet distributed to members and which

- 13 -

can be subscribed to. The first issues will be largely dedicated to the publication of descriptive memoranda on works and articles concerning the work of the ATALA, the bulletin thus being able to go back to and continue the bibliographic work undertaken since 1958 within the international group. (Bibliographie de la traduction automatique [Bibliography of machine translation], by E. and K. Delavenay, The Hague, Mouton and Co. (to be published).)

Material organization of the work.

The provisional office of the association has until now abstained from taking on important financial engagements, desiring to establish social resources solidly before doing that. However, it is permitted to consider the establishment in the first half of 1960 of a small permanent nucleus, assuring the secretariat and necessary liaisons, and furnishing a base of operations for the benevolent works of members. This question could probably be settled after the establishment, by the first annual general Assembly, of the council of administration stipulated by the statutes. It already appears indispensable to consider the appointment of several researchers working full time and able to support and consolidate the work of the scientific committees. It is not out of place to recall in this connection that the first contributions of fifteen or sixteen member-benefactors would permit us to pay for the services of a young researcher on a regular basis.

Theoretical works.

The very composition of the Association assures for itself the possibility of stimulating among its members theoretical works of exploration in fields bordering on applied linguistics. Until now works of two orders have been retained as capable of illuminating certain important aspects of practical work. On the one hand a study is stipulated, directly connected with the translation of English into French, on the comparative stylistics of scientific English and French. It will be a matter of applying to stylistics the methods already perfected by Mr. J. C. Gardin for archaeological studies, and of mapping out with the aid of precise measuring instruments the limit between the strictly linguistic processes of expression and those which are within the province of stylistics. On the other hand, a scientific committee charged with digging up works already undertaken or which are possible on languages other than English and French, thinks that it can turn with advantage to the problem of the "congruence" of languages, which touches closely that of the pivot language. By congruence of two languages we understand the degree of facility with which they can be translated one into the other, which is not necessarily connected with their resemblance or their common origins.

These projects of approach toward the problems of applied linguistics, as moreover the number and quality of the members of the new Association, show that French scholars and linguists are ready to

- 14 -

act together as soon as the cadre for their collective action is established. It is no longer a question here in effect of individual tasks carried out over a lifetime by isolated individuals. Applied linguistics brings into play techniques and knowledge so diverse and requires so onerous a set of tools, that only interdisciplinary groups established around collective means of action can actively participate in it. On different sides initiatives are, and will be, taken to endow France in this field with the means of research in conformity with her needs and her possibilities. If, as must be hoped, full time projects will soon be undertaken to implement machine translation programs of scientific texts, the ATALA will be in a position to stimulate the interest of scholars and young people in this research. With official laboratories it will play a role of information and liaison; it will be able to undertake limited and precise research on particular aspects of certain problems; and finally it will freely assure scientific liaison with centers of foreign research. That is a modest role compared with the breadth of problems which are posed in connection with language as a means of communication. But we must not under-estimate the possible range of a limited and precise action, carried out with tenacity with a view to making our country participate in a new scientific movement. The Association which has just been founded has already permitted the grouping of scattered initiatives and of tendencies which were mutually unaware of each other. It will be able to remain in the forefront with respect to official research and to orient the latter when required towards new problems and dynamic ideas.

<div align="center">Emile Delavenay.</div>

5299   - END -